

IN THE CLAIMS:

Please cancel Claims 98-121 without prejudice to or disclaimer of the subject matter contained therein.

Please add new Claims 122-170 as follows.

1-121. (Cancelled)

122. (New) An apparatus for generating annotation data for use in annotating a data file comprising audio data, the apparatus comprising:

an automatic speech recognizer operable to generate phoneme data for audio data in the data file;

a word decoder operable to identify words within the phoneme data generated by the automatic speech recognizer; and

an annotation generator operable to generate annotation data by combining the generated phoneme data and the words identified by said word decoder.

123. (New) An apparatus for generating annotation data for use in annotating a data file, the apparatus comprising:

an input operable to receive an input voice signal;

a speech recognizer operable to convert the input voice signal into phoneme data and words; and

an annotation generator operable to generate annotation data by combining the phoneme data and the words.

124. (New) An apparatus for generating annotation data for use in annotating a data file, the apparatus comprising:

an input operable to receive image data representative of text;

a character recognizer operable to convert said image data into text data;

a converter operable to convert words in the text data into phoneme data; and

an annotation generator operable to generate annotation data by combining the phoneme data and words in the text data.

c/ 125. (New) An apparatus according to claim 122, wherein said annotation data defines a phoneme and word lattice and wherein said annotation generator comprises:

(i) a first generator operable to generate data defining a plurality of nodes within the lattice and a plurality of links connecting the nodes within the lattice;

(ii) a second generator operable to generate data associating a plurality of phonemes of the phoneme data with a respective plurality of links; and

(iii) a third generator operable to generate data associating at least one of the words with at least one of said links.

126. (New) An apparatus according to claim 125, wherein said annotation generator is operable to generate said data defining said phoneme and word lattice in blocks of said nodes.

127. (New) An apparatus according to claim 125, wherein said annotation generator is operable to generate data defining time stamp information for each of said nodes.

128. (New) An apparatus according to claim 127, wherein said annotation generator is operable to generate said phoneme and word lattice data in blocks of equal time duration.

21 129. (New) An apparatus according to claim 126, wherein said annotation generator is operable to generate data which defines each block's location within a database.

130. (New) An apparatus according to claim 127, wherein said data file includes a time sequential signal, and wherein said annotation generator is operable to generate time stamp data which is time synchronized with said time sequential signal.

131. (New) An apparatus according to claim 122, wherein said audio data includes audio data which defines the speech of a plurality of speakers, and wherein said annotation generator is operable to generate data which defines separate phoneme and word annotation data for the speech of each speaker.

132. (New) An apparatus according to claim 122, wherein said speech recognizer is operable to generate data defining a weighting for the phonemes in the phoneme data.

133. (New) An apparatus according to claim 122, wherein said word decoder is operable to generate data defining a weighting for the words identified within said phoneme data.

134. (New) An apparatus according to claim 125, wherein said first generator is operable to define at least one node which is connected to a plurality of other nodes by a plurality of links.

C/ 135. (New) An apparatus according to claim 123, wherein said speech recognizer is operable to generate data defining a weighting for the phonemes in the phoneme data.

136. (New) An apparatus according to claim 135, wherein said speech recognizer is operable to generate data defining a weighting for the words.

137. (New) An apparatus according to claim 123, further comprising an associator operable to associate said annotation data with said data file.

138. (New) An apparatus according to claim 124, wherein said converter comprises an automatic phonetic transcription unit which generates said phoneme data from words within the text data output by said character recognizer.

139. (New) An apparatus according to claim 124, further comprising an associator operable to associate said annotation data with either said image data representative of said text or with said text data.

140. (New) An apparatus according to claim 124, wherein said input comprises a document scanner or a facsimile machine.

C1 141. (New) A method of generating annotation data for use in annotating a data file comprising audio data, the method comprising the steps of:

using an automatic speech recognizer to generate phoneme data for audio data in the data file;

using a word decoder to identify words within the phoneme data generated by the automatic speech recognizer; and

generating annotation data by combining the generated phoneme data and the words identified by said word decoder.

142. (New) A method of generating annotation data for use in annotating a data file, the method comprising the steps of:

receiving an input voice signal;
processing the input voice signal using a speech recognizer to generate phoneme data and word data for the input voice signal; and
generating annotation data by combining the phoneme data and the word data generated for the input voice signal.

143. (New) A method of generating annotation data for use in annotating a data file, the method comprising the steps of:

receiving image data representative of text;
converting said image data into text data using a character recognizer;
converting words in the text data into phoneme data; and
generating annotation data by combining the phoneme data and words within the text data.

144. (New) A method according to claim 141, wherein said annotation data defines a phoneme and word lattice and wherein said generating step comprises the steps of:

(i) generating data defining a plurality of nodes within the lattice and a plurality of links connecting the nodes within the lattice;
(ii) generating data associating a plurality of phonemes of the phoneme data with a respective plurality of links; and
(iii) generating data associating at least one of the words with at least one of said links.

145. (New) A method according to claim 144, wherein said generating step generates data defining time stamp information for each of said nodes.

146. (New) A method according to claim 145, wherein said generating step generates data which defines each block's location within a database.

147. (New) A method according to claim 145, wherein said data file includes a time sequential signal, and wherein said generating step generates time stamp data which is time synchronized with said time sequential signal.

148. (New) A method according to claim 141, wherein said audio data includes audio data which defines the speech of a plurality of speakers, and wherein said generating step generates data which defines separate phoneme and word annotation data for the speech of each speaker.

149. (New) A method according to claim 141, wherein said speech recognizer generates data defining a weighting for the phonemes associated with said links.

150. (New) A method according to claim 141, wherein said word decoder generates data defining a weighting for the words associated with said links.

151. (New) A method according to claim 144, wherein said step of defining a plurality of nodes and a plurality of links defines at least one node which is connected to a plurality of other nodes by a plurality of links.

152. (New) A method according to claim 142, wherein said speech recognizer generates data defining a weighting for the phonemes associated with said links.

153. (New) A method according to claim 142, wherein said speech recognizer generates data defining a weighting for the words associated with said links.

21 154. (New) A method according to claim 142, further comprising the step of associating said annotation data with said data file.

155. (New) A method according to claim 143, wherein said step of converting words into phoneme data uses an automatic phonetic transcription unit which generates said phoneme data for words within the text data output by said character recognizer.

156. (New) A method according to claim 143, further comprising the step of associating said annotation data with either said received image data or with said text data.

157. (New) A method according to claim 143, wherein said receiving step uses a document scanner or a facsimile machine.

158. (New) An apparatus according to claim 122, wherein said annotation generator is operable to generate annotation data that defines a phoneme and word lattice arranged in a time-ordered sequence of blocks.

159. (New) An apparatus according to claim 158, wherein said phoneme and word lattice is associated with a time-sequential signal, and wherein said annotation generator is operable to generate said phoneme and word lattice so that:

(i) the time-ordered sequence of blocks of the phoneme and word lattice is time-synchronized with the time-sequential signal; and

(ii) each block of the phoneme and word lattice includes an associated time index identifying a timing of the block within the time-sequential signal.

160. (New) An apparatus according to claim 159, wherein said annotation generator is operable to generate node data which defines a plurality of nodes within the lattice, each node representing a point in time at which a word and/or phoneme begins or ends within the associated time-sequential signal and wherein each node includes a time-offset value defining the point in time represented by the node relative to said time index associated with the corresponding block.

161. (New) An apparatus according to claim 159, wherein said annotation generator is operable to generate node data which defines a plurality of nodes within the lattice,

each node representing a point in time at which a word and/or phoneme begins or ends within the associated time-sequential signal and wherein each node includes a time-offset value defining the point in time represented by the node relative to said time index associated with the corresponding block.

162. (New) An apparatus according to claim 122, wherein said automatic speech recognizer is operable to generate phoneme data that includes a phoneme lattice which identifies a number of different possible phoneme strings which correspond to speech within the audio data, and wherein said annotation generator is operable to generate said annotation data by combining the generated phoneme lattice with the words identified by the word decoder.

163. (New) An apparatus according to claim 123, wherein said automatic speech recognizer is operable to generate phoneme data that includes a phoneme lattice which identifies a number of different possible phoneme strings which correspond to speech within the audio data, and wherein said annotation generator is operable to generate said annotation data by combining the generated phoneme lattice with the words identified by the word decoder.

164. (New) An apparatus according to claim 122, wherein said annotation generator is operable to generate header data relating to the speech recognizer that generated the phoneme data for the annotation.

165. (New) An apparatus according to claim 164, when said annotation generator is operable to generate header data including data identifying the language and the phoneme sets used by the automatic speech recognizer.

166. (New) An apparatus according to claim 123, wherein said annotation generator is operable to generate header data relating to the speech recognizer that generated the phoneme data for the annotation.

167. (New) An apparatus according to claim 166, wherein said annotation generator is operable to generate header data including data identifying the language and the phoneme sets used by the automatic speech recognizer.

168. (New) A computer program code for causing a programmable processing apparatus to become operable to generate annotation data for use in annotating a data file comprising audio data, the code comprising:

code for causing an automatic speech recognizer to generate phoneme data for audio data in the data file;

code for causing a word decoder to identify words within the phoneme data represented by the automatic speech recognizer; and

code for generating annotation data by combining the generated phoneme data and the words identified by the word decoder.

169. (New) A computer program code according to claim 167, carried by a carrier signal.

170. (New) A computer readable medium storing computer executable instructions for causing a programmable processing apparatus to become operable to generate annotation data for use in annotating a data file comprising audio data, the medium comprising:

code for causing an automatic speech recognizer to generate phoneme data for audio data in the data file;

code for causing a word decoder to identify possible words within the phoneme data represented by the automatic speech recognizer; and

code for generating annotation data by combining the generated phoneme data and the words identified by said word decoder.
